

INTERNATIONAL STANDARD

ISO/IEC
11179-5

First edition
1995-12-01

Information technology — Specification and standardization of data elements —

Part 5:

Naming and identification principles for data
elements

*Technologies de l'information — Spécifications et normalisation des
données —*

Partie 5: Principes de dénomination et d'identification des données



Reference number
ISO/IEC 11179-5:1995(E)

Contents

Foreword.....	iii
Introduction.....	iv
1. Scope.....	1
2. Normative references	1
3. Definitions	1
4. Principles for the identification structure of data.....	3
4.1 Identifying attributes	3
4.2 Name and context	3
4.3 International registration data identifier	4
5. Rules for registration identification of data	4
6. Guidelines for structured naming conventions.....	5
6.1 Principles governing semantic content of names	5
6.2 Principles governing format of names	9
7. Thesaurus application guidelines	9
Annexes	
A Example naming convention	10
B Registration identification example	12
C Thesaurus example.....	14

© ISO/IEC 1995

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

ISO/IEC Copyright Office • Case postale 56 • CH-1211 Genève 20 • Switzerland
Printed in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC1. Draft international Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75% of the national bodies casting a vote.

International Standard ISO/IEC 11179-5 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, SC 14, *Data element principles*.

ISO/IEC 11179 consists of the following parts, under the general title *Information technology - Specification and standardization of data elements*:

- *Part 1: Framework for the generation and standardization of data elements*
- *Part 2: Classification of concepts for the identification of domains*
- *Part 3: Basic attributes of data elements*
- *Part 4: Rules and guidelines for the formulation of data definitions*
- *Part 5: Naming and identification principles for data elements*
- *Part 6: Registration of data elements*

Annexes A, B and C of this part of ISO/IEC 11179 are for information only.

Introduction

This part of ISO/IEC 11179 contains principles, rules and guidelines. Principles establish the premises on which the rules are based. Rules are mandatory and testable for compliance. Guidelines are applications of the rules recommended for good practice.

Information technology - Specification and standardization of data elements -

Part 5: Naming and identification principles for data elements

1 Scope

This part of ISO/IEC 11179 provides rules and guidelines for naming and identification of data elements. It describes the components and structure of data element identification. Identification is narrowly defined to encompass only the means to establish unique identification of data elements within a register. It defines the identifying attributes; describes the relationship of the attributes to each other; includes principles by which naming conventions can be developed; and describes an example naming convention. The naming guidelines described herein can also be applied to names of data concepts in addition to names of data elements.

This part of ISO/IEC 11179 should be used in conjunction with those which establish rules and procedures for attributing, classifying, defining, and registering data elements.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO/IEC 11179. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO/IEC 11179 are encouraged to apply the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO/IEC 11179-3:1994, *Information technology - Specification and standardization of data elements - Part 3: Basic attributes of data elements*.

ISO/IEC 11179-4:1995, *Information technology - Specification and standardization of data elements - Part 4: Rules and guidelines for the formulation of data definitions*

3 Definitions

For the purposes of this part of ISO/IEC 11179, the following definitions apply.

3.1 attribute: A characteristic of an object or entity (ISO/IEC 11179-3).

3.2 context: A designation or description of the application environment or discipline in which a name is applied or from which it originates (ISO/IEC 11179-3).

3.3 data element: A unit of data for which the identification, meaning, representation and permissible values are specified by means of a set of attributes (ISO/IEC 11179-3).

3.4 data element concept: A concept which can be represented in the form of a data element, described independently of any particular representation (ISO/IEC 11179-3).

3.5 data identifier: An identifier of a data element (a string of characters or other graphic symbols) assigned by a Registration Authority.

3.6 definition: A word or phrase expressing the essential nature of a person or thing or class of persons or things: an answer to the question "what is x?" or "what is an x?"; a statement of the meaning of a word or word group (Webster's Third New International Dictionary of the English Language Unabridged, 1986).

3.7 identifier: See data identifier.

3.8 lexical: Pertaining to words or the vocabulary of a language as distinguished from its grammar and construction.

3.9 name: The primary means of identification of objects and concepts for humans (ISO/IEC 11179-4).

3.10 object class term: A component of the name of a data element which represents the logical data grouping (in a logical data model) to which it belongs; e.g., "employee."

3.11 property term: A component of the name of a data element which expresses the category to which the data element belongs.

3.12 qualifier term: A word or words which help define and differentiate a name within the database.

3.13 registration authority: An organization authorized to register a data element.

3.14 registration authority identifier: The identifier of any organization authorized to register a data element; the attribute which stores this.

3.15 representation term: The form of the set of valid values for a data element, e.g., "amount," "name."

3.16 semantics: The branch of linguistic science which deals with the meanings of words (Webster).

3.17 separator: A symbol or space enclosing or separating a component within a name; a delimiter.

3.18 structure set: A method of placing objects in context, revealing relationships to other objects. Examples include Entity-Relationship Models, taxonomies, and ontologies.

3.19 syntax: The relationships among characters or groups of characters, independent of their meanings or the manner of their interpretation and use. The structure of expressions in a language, and the rules governing the structure of a language.

3.20 thesaurus: A controlled vocabulary arranged in a given order in which relationships among terms are displayed and identified.

3.21 version identifier: The identification of an issue of a data element specification in a series of evolving data element specifications within a Registration Authority.

4 Principles for the identification structure of data

4.1 Identifying attributes

A set of five related attributes serves to name and identify each data element for the purpose of differentiating data elements. These attributes are:

- name
- context
- registration authority identifier
- data identifier
- version identifier

References to name in this part of ISO/IEC 11179 shall be assumed to include both the name and synonymous name attributes specified in ISO/IEC 11179-3, which contains complete descriptions of all attributes. Also, data identifier is equivalent to identifier in ISO/IEC 11179-3. The principles for these attributes are stated below.

4.2 Name and context

A data element shall have at least one name within a register of a Registration Authority. More names may be assigned depending on the context in which the data element is used. Each name has special utility within a particular context. Rigorously structured names may be created for data administration, a preferred name may be specified by users, shortened names may be generated for particular software environments such as a particular programming language or database management system.

Within each context names for many data elements may be assigned. A naming convention (usually a set of rules) is established for each context to specify how names are formulated within that context. A naming convention shall cover all pertinent aspects of the context. This includes, as applicable:

- the scope of the naming convention, e.g., industry preferred name;
- the authority which establishes names;
- semantic rules governing the source and content of the words used in a name, e.g., words derived from data models, words commonly used in the discipline, etc.;
- syntactic rules covering required word order;
- lexical rules covering controlled word lists, name length, character set, language;
- a rule establishing whether or not names within this context must be unique.

These aspects of a naming convention are detailed in Clause 7, which provides guidelines for developing a rigorously structured naming convention.

4.3 International registration data identifier

The attributes registration authority identifier (RAI), data identifier (DI), and version identifier (VI) constitute the international registration data identifier (IRDI). At least one IRDI is required for a data element. Data identifiers are assigned by a Registration Authority; data identifiers shall be unique within a domain of a Registration Authority.

As each Registration Authority may determine its own assignment scheme, there is no guarantee that the DI by itself will uniquely identify a data element. For example, if two authorities both use sequential 6-digit numbers, there will be a set of data elements with the same DI's; however, in all probability, the data elements with the same numbers will not be the same data elements. Conversely, if one data element appears in two registers, it will have two DI's. Therefore, both the DI and the RAI are necessary for identification of a data element.

If particular attributes of a data element change, then a new version of the data element shall be created and registered. In such a case, a VI is required to complete the unique identification of a data element. For further guidance, see ISO/IEC 11179-6.

An IRDI can serve as a key when exchanging data among information systems, organizations, or other parties who wish to share a specific data element, but may not utilize the same names or contexts. An IRDI is also useful for language translation when the IRDI is associated with contexts established for more than one natural language and referencing among sets of data elements controlled by different Registration Authorities.

ISO/IEC 11179 does not specify the format or content of a unique DI.

Requirements for a Registration Authority, and a discussion of the IRDI, appear in ISO/IEC 11179-6.

5 Rules for registration identification of data

- 1) Each data element shall have a unique data identifier within the register of a Registration Authority.
- 2) The combination of registration authority identifier, data identifier, and version identifier shall constitute a unique identification of a data element.

3) To be assigned a data identifier, a data element shall be:

- derived according to ISO/IEC 11179-2,
- attributed according to ISO/IEC 11179-3,
- defined according to ISO/IEC 11179-4,
- named according to ISO/IEC 11179-5, and
- registered according to ISO/IEC 11179-6.

4) A data element shall have at least one name within a context.

6 Guidelines for structured naming conventions

The following are guidelines that could be used to develop a naming convention to produce rigorously structured names for a particular context. Annex A is an example of a specific naming convention that is consistent with the guidelines presented in this Clause.

The guidelines are described in general terms with examples furnished. Rules are derived from the principles by which names are developed; these rules form a naming convention. Names formed according to these rules can be easily translated into languages other than the original because of the simplified syntax. Syntactic, semantic and lexical rules vary by organizations such as corporations or standards-setting bodies for business sectors; each can establish rules for name formation within its context.

As discussed in subclause 6.1.1.1, each data element is formed from a set of components selected from the structure sets within its context.

Data element names can be formed from the names of components, each assigned meaning (semantics) and relative or absolute position (syntax) within a name. They may be subject to lexical rules. They may, but need not, be delimited by a separator symbol. The set or range of values of each component should be rigorously controlled by an authority, e.g., a data administrator within a corporation or an approving committee for an international business sector naming standard.

- Semantic rules enable meaning to be conveyed;
- Syntactic rules relate components in a consistent, specified order;
- Lexical (word form and vocabulary) rules reduce redundancy and increase precision;

6.1 Principles governing semantic content of names

Semantics concerns the meanings of name components and possibly separators which delimit them.

6.1.1 Semantics of name components

Components consist of discrete terms. The components described in this part of ISO/IEC 11179 are:

- object class terms;
- property terms;
- representation terms;
- qualifier terms;

6.1.1.1 Object class term

An object class term is a component of a data element name which represents an activity or object in a context. Using a modelling methodology, as for instance an Entity Relationship Diagram (ERD) or object model, is a way to locate and discretely place all data elements in relation to their higher-level model entities. The attributes of entity-relationship model entities equate to data elements which are related to each other through further application of the methodology. In an object model, data elements are expressed as object attributes.

Models provide one kind of classification scheme for data elements. Data elements may be identified with their related modelling entities by mapping the object class term to the model entity name. ISO/IEC 11179-1 provides examples of the mapping between object class terms and ERD and object model entities.

For example, in the data elements:

- Employee Last Name
- Cost Budget Period Total Amount
- Tree Height Measure
- Member Last Name

the components Employee, Cost, Tree, and Member are object class terms.

6.1.1.2 Property term

A set of property terms is composed from a set of name components in a property taxonomy. This set must consist of terms which are discrete (the definition of each does not overlap the definition of any other), and complete (taken together, the set represents all information concepts required for the specification of data elements).

For example, in the data elements:

- Employee Last Name
- Cost Budget Period Total Amount
- Member Last Name
- Tree Height Measure

the components Last Name, Total Amount, and Height are properties.

The property term will occur naturally in the definition of a data element.

Using components from two structure sets provides a complementary way of categorization. Both object class and property components of data elements are utilized to form a name which contains vital information about the data element, and also excludes extraneous or irrational elements which may be introduced when no conventions are employed.

6.1.1.3 Representation term

A representation term is a component of a data element name which describes the form of representation of the data element. Each term is developed from a controlled word list or a taxonomy. Representation terms categorize forms of representation such as:

- | | | |
|------------|----------|-------|
| • Name | • Amount | |
| • Measure | • Number | • • • |
| • Quantity | • Text | |

This term describes the form of the set of valid values of a data element. Often, the representation term may be redundant with part of the property term. When this occurs, one term or part of one term may be eliminated in a structured name. This can be established as a rule in a naming convention. See Annex A for an application of this procedure.

For example, in the data elements:

- Tree Height Measure
- Employee Last Name

the components Measure and Name are representation terms. Note that Last Name is a property term. One occurrence of the redundant word Name is removed to promote clarity.

6.1.1.4 Qualifier term

Qualifier terms may be attached to object class terms, property terms, and representation terms if necessary to uniquely identify a data element. These qualifier terms may be derived from structure sets specific to a context. In the rules for a naming convention, a restriction in the number of qualifier terms is recommended.

For example, in the data element:

- Cost Budget Period Total Amount

the component Budget Period is a qualifier term.

NOTE – Limitations in the form of permitted terms of qualifiers help reduce redundancy and increase incidence of data reuse by eliminating synonyms. This applies also to object class terms, property terms, and representation terms. A mechanism such as a thesaurus of terms facilitates this effort (see Clause 7; Annex C).

6.1.2 Semantics of separators

Name components are delimited by separators. These may have:

- No semantic meaning. A naming rule may state that separators will consist of one blank space or exactly one special character (for example a hyphen or underscore) regardless of semantic relationships of components. Such a rule simplifies name formation.
- Semantic meaning. Semantic meaning can be conveyed by separators by, for example, assigning a different separator between words in the qualifier term from the separator which separates words in the other component terms. In this way, the separator identifies the qualifier term clearly as different from the rest of the name.

For example, in the data element:

- Cost-Budget_Period-Total-Amount

the separator between words in the qualifier term is an underscore; other name components are separated by hyphens.

Some languages, such as German and Dutch, commonly join grammatical constructs together in a single word (resulting in one word which in English or French might be a phrase consisting of nouns and adjectives). These languages could use a separator which is not a break between words, such as a hyphen, space or underscore, but instead capitalize the first letter of each name component within a single word.

6.2 Principles governing format of names

6.2.1 Syntactic principles

Syntactic principles specify the arrangement of components within a name. This arrangement may be specified as relative or absolute, or some combination of the two.

- 1) Relative arrangement specifies components in terms of other components, e.g., a rule within a convention might require that a qualifier must always appear before the component being qualified appears.
- 2) Absolute arrangement specifies a fixed occurrence of the component, e.g., a rule might require that the property is always the last component in a name.

6.2.2 Lexical principles

These principles concern preferred and non-preferred terms, synonyms, abbreviations, component length, spelling, permissible character set, case sensitivity, etc.

7 Thesaurus application guidelines

A thesaurus in which the user can find a variety of synonyms, near-synonyms and homographs for name components is a valuable tool. It can provide semantic linking between preferred name terms and other terms. In addition to guidance for use of homographs (words with the same spelling representing different concepts), a thesaurus can direct the user through choices involving:

- **equivalence** - in which one word or term is preferred over others for expression of a concept;
- **hierarchy** - in which a relationship between broader and narrower terms is expressed by levels of superordination or subordination; and
- **association** - in which two or more terms are semantically or conceptually associated, whether they belong to the same hierarchy or different hierarchies.

A Registration Authority may develop a thesaurus for name components and distribute it to interested parties; in addition, development of subject area thesauri is recommended.

A thesaurus may be used to describe structures in verbal representation as a supplement to graphic depiction. Preferred terms become name components; relationships between preferred terms express the position of components in the structure.

A controlled vocabulary is an advantage for thesaurus use. Control can be built into a thesaurus through scope of descriptors, linking of synonyms and near-synonyms through equivalence, and the resolution of homographs. These are all functions which users need to navigate through the system.

Annex A

(informative)

Example naming convention

These rules are derived from the guidelines described in Clause 6. Examples are included. They may be applied to the development of context names at the discretion of the subject area authority.

A.1 Semantic rules

- a) Object classes represent things of interest in a universe of discourse which may, for instance, be found in a model of that universe.

Example: **Cost**

- b) One and only one object class term shall be present.

- c) Property terms shall be derived from the property system structure set and represent the category of the data.

Example: **Total Amount**

- d) One and only one property term shall be present.

- e) Qualifiers may be derived as determined by the subject area authority and will be added as needed to describe the data element and make it unique within a specified context. The order of the qualifier terms is not significant. Qualifier terms are optional.

Example: **Budget Period**

- f) The representation of the valid value set of the data element is described by the representation term.

- g) One and only one representation term shall be present.

Example: **Amount**

A.2 Syntactic rules

- a) The object class term shall occupy the first (leftmost) position in the name.
- b) Qualifier terms shall precede the component qualified. The order of qualifiers shall not be used to differentiate data element names.
- c) The property term shall occupy the next position.

- d) The representation term shall occupy the last position. If any word in the representation term is redundant with any word in the property term, one occurrence will be deleted.

Example: **Cost Budget Period Total Amount**

A.3 Lexical rules

- a) Nouns are used in singular form only. Verbs (if any) are in the present tense.
- b) Name components and words in multi-word terms are separated by spaces. No special characters are allowed.
- c) All words in the name are in mixed case.
- d) Abbreviations, acronyms, and initialisms are allowed.

Example: **Cost Budget Period Total Amount**

A.4 Uniqueness rule

All names shall be unique within this context.

Annex B

(informative)

Registration identification example

(English language)

The identification structure of an example data element is as follows:

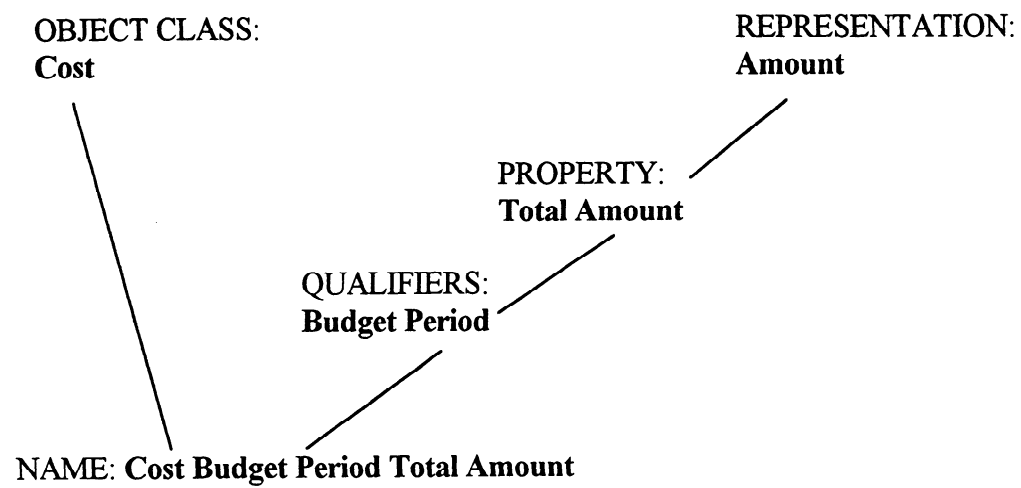
DATA ELEMENT

contains

RAI	DI	VERSION	NAME	CONTEXT
ISO	848575	1	Cost Budget Period Total Amount	FIN-EDI System
ISO	848575	1	ACCOUNT_AMOUNT	USA_GICS System
IEEE	193847	1	Transfer-Cost-Amount	Engineering Model
IEEE	193847	1	our_cost_\$	Contracts Report Header

This is an example of an element which is registered by two Registration Authorities. Each authority assigns its own data identifier. Under each authority, the element has two different names. Note that ACCOUNT_AMOUNT and our_cost_\$ are examples of synonymous names in application systems. These legacy names are those often created without recourse to naming conventions, and then perpetuated over time.

The component structure for this element (according to the guidelines of Clause 6) is as follows:



Annex C

(informative)

Thesaurus example

An excerpt from a possible thesaurus of structure terms, including structural information as well as synonym and homonym resolution.

key:

BT - Broader Term

NT - Narrower Term

UF - Use For

USE - Use the following term instead

RT - Related Term

SN - Scope Note

THESAURUS EXCERPT

COST

SN Amount the organization spends to procure goods or services.

BT Contract

NT Petty Cash

UF *Expense*

RT Budget

RT Amount

Expense

USE Cost

Note that although the thesaurus entry can show that COST is related to both Budget and Amount, the exact nature of the relationship is not explicit.

Italics are used to denote deprecated terms. This is the method used to distinguish preferred terms among synonyms.

One or more levels of hierarchy can be shown by listing several broader and narrower terms. This is a thesaurus design decision.

The scope note (SN) reflects the definition as described in ISO/IEC 11179-4.

This page intentionally left blank

ICS 35.040

Descriptors: data processing, information interchange, data representation, data elements, specifications, identification methods, naming.

Price based on 14 pages
